
Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPSrch_pp protein - protein database search, using Smith-Waterman algorithm
Run on: Wed Apr 12 07:31:27 2000; MasPar time 8.98 Seconds
Tabular output not generated.
578.952 Million cell updates/sec
Title: >US-09-276-268-12
Description: (1-174) from US09276268.pep
Perfect Score: 1353
Sequence: 1 MAPANLGLTPHVMVLGAVL.....HDEIRKKGFLKQNPYKRF 174
Scoring table: PAM 150
Gap 11
Searched: 82229 seqs, 29864866 residues
Post-processing: Minimum Match 0%
Listing first 45 summaries
Database: swiss-prot38
1:swissprot
Statistics: Mean 44.875; Variance 85.350; scale 0.526
Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.
SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description	Pred. No.
1	1065	78.7	180	1	C211_HUMAN PUTATIVE SURFACE GLYCO	9.67e-199
2	143	10.6	319	1	A33_HUMAN CELL SURFACE A33 ANTIG	8.83e-08
3	129	9.5	1898	1	TRHY_HUMAN TRICHOHYALIN.	1.42e-05
4	118	8.7	472	1	YWIE_CAEEL HYPOTHETICAL 54.3 KD P	6.43e-04
5	113	8.4	463	1	CD4_CANFA T-CELL SURFACE GLYCOP	3.43e-03
6	114	8.4	471	1	RU17_XENLA U1 SMALL NUCLEAR RIBON	2.46e-03
7	114	8.4	815	1	ERKS_HUMAN EXTRACELLULAR SIGNAL-R	2.46e-03
8	112	8.3	488	1	UZAF_CAEEL SPLICING FACTOR UZAF 6	4.77e-03
9	111	8.2	503	1	YOW5_CAEEL HYPOTHETICAL 57.1 KD P	6.63e-03
10	111	8.2	793	1	CALD_HUMAN CALDESMON (CDM).	6.63e-03
11	110	8.1	134	1	CLX2_HUMAN COMPLEXIN 2 (SYNAPHIN	9.18e-03
12	110	8.1	614	1	RU17_HUMAN U1 SMALL NUCLEAR RIBON	9.18e-03
13	108	8.0	422	1	SYT2_RAT SYNAPTOTAGMIN II.	1.76e-02
14	108	8.0	422	1	SYT2_MOUSE SYNAPTOTAGMIN II.	1.76e-02
15	108	8.0	531	1	CALD_RAT NON-MUSCLE CALDESMON (1.76e-02
16	108	8.0	1407	1	TRHY_RABIT TRICHOHYALIN.	1.76e-02
17	107	7.9	361	1	EB12_HUMAN EBV-INDUCED G PROTEIN-	2.42e-02
18	105	7.8	519	1	YMP4_YEAST HYPOTHETICAL 60.1 KD P	4.58e-02
19	105	7.8	1085	1	YAF4_SCHPO HYPOTHETICAL 122.9 KD	4.58e-02
20	105	7.8	2468	1	MAPB_HUMAN MICROFUBULE-ASSOCIATED	4.58e-02
21	104	7.7	244	1	YL53_CAEEL HYPOTHETICAL 29.0 KD P	6.28e-02
22	104	7.7	769	1	ITB2_BOVIN CELL SURFACE ADHESION	6.28e-02
23	103	7.6	59	1	HSPI1_MACRU SPERM PROTAMINE P1.	8.59e-02

24	102	7.5	134	1	CLX2_NARJA COMPLEXIN 2 (NJ-SYNAPH	1.17e-01
25	101	7.5	426	1	SY65_APLCA SYNAPTOTAGMIN (P65).	1.60e-01
26	101	7.5	432	1	GFAP_HUMAN GLIAL FIBRILLARY ACIDI	1.60e-01
27	101	7.5	507	1	VGLG_VHSVO SPIKE GLYCOPROTEIN PRE	1.60e-01
28	101	7.5	771	1	ITB2_MOUSE CELL SURFACE ADHESION	1.60e-01
29	102	7.5	880	1	DL_DROME NEUROGENIC LOCUS DELTA	1.17e-01
30	101	7.5	1150	1	YKDI_CAEEL HYPOTHETICAL 132.3 KD	1.60e-01
31	102	7.5	1200	1	DDX8_CAEEL PUTATIVE PRE-mRNA SPLI	1.60e-01
32	101	7.5	1240	1	YNU1_YEAST HYPOTHETICAL 141.5 KD	1.60e-01
33	100	7.4	239	1	CALD_MELGA CALDESMON, SMOOTH MUSC	2.18e-01
34	100	7.4	241	1	YOB7_CAEEL HYPOTHETICAL 28.5 KD P	2.18e-01
35	100	7.4	434	1	YK12_YEAST HYPOTHETICAL 50.5 KD P	2.18e-01
36	100	7.4	650	1	Y096_MYCGE HYPOTHETICAL PROTEIN M	2.18e-01
37	100	7.4	771	1	CALD_CHICK CALDESMON (CDM).	2.18e-01
38	99	7.3	444	1	CEB_DROME CCAAT/ENHANCER BINDING	2.95e-01
39	98	7.2	178	1	BTC_HUMAN BETACELLULIN PRECURSOR	4.00e-01
40	97	7.2	235	1	BAR6_CHITE BALBIANI RING PROTEIN	5.41e-01
41	98	7.2	330	1	RLX3_STAUD RLX PROTEIN.	4.00e-01
42	98	7.2	473	1	TED_HUMAN MNM4 PROTEIN.	5.41e-01
43	97	7.2	1178	1	DDX8_HUMAN PROBABLE ATP-DEPENDENT	5.41e-01
44	97	7.2	1220	1	YOD2_CAEEL HYPOTHETICAL 26.8 KD P	7.29e-01
45	96	7.1	233	1		

ALIGNMENTS

RESULT 1
ID C211_HUMAN STANDARD; PRT; 180 AA.
AC P53801;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE PUTATIVE SURFACE GLYCOPROTEIN C210RF1 PRECURSOR (C210RF3).
GN C210RF1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia;
OC Eutheria; Primates; Catarrhini; Homnidae; Homo.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=THYMUS;
RX MEDLINE; 98234553.
RA YASPO M.-L., AALTONEN J., HORELLI-KUITONEN N., PELTONEN L.,
RA LEHRACH H.;
RT *Cloning of a novel human putative type Ia integral membrane protein
mapping to 21q22.3.*;
RL Genomics 49:133-136(1998).
CC -1- FUNCTION: MAY PLAY AN IMPORTANT ROLE IN A CELL TRAFFICKING
MECHANISM.
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (POTENTIAL).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC EMBL; Z50022; CAA90325.1;
DR MIM; 603784;
KW Transmembrane; Glycoprotein; Signal.
FT SIGNAL 1 32 POTENTIAL.
FT CHAIN 33 180 PUTATIVE SURFACE GLYCOPROTEIN C210RF1.
FT DOMAIN 33 96 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 97 117 POTENTIAL.
FT DOMAIN 118 180 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 116 121 POLY-CYS
FT CARBOHYD 45 45 POTENTIAL.
FT CARBOHYD 54 54 POTENTIAL.
SQ SEQUENCE 180 AA; 20323 MW; 8424FF4D CRC32;
Query Match 78.7% Score 1065; DB 1; Length 180;
Best Local Similarity 75.1%; Pred. No. 9.67e-199;

Matches 133; Conservative 25; Mismatches 16; Indels 3; Gaps 3;

Db 1 MAPVARGPTPYRRLGGNALLLIPVAAQEPGAACQNTNTCECLKNVSLWC 60
 QY 1 MAPA-NLGTPTPHVMLG-AVLLLL-SGASQAPPRVGCSEYTNCSCECLRNVSCLWC 57
 Db 61 NTNACLDYPTVSVLPASICKLSSARWGVNFWNFEALITMSVLGSLGTVCCVC 120
 QY 58 NENKACMDYPTVSVLPASICKLSSARWGVNFWNFEALITMSVLGSLGTVCCVC 117
 Db 121 CRKRSRKRDRSEKERRERERRRQERRAEMKTRHDEIRKKYGLFKEENPYARP 177
 QY 118 CRKRSRKRDRSEKERRERERRRQERRAEMKTRHDEIRKKYGLFKEENPYARP 174

RESULT 2

ID A33_HUMAN STANDARD; PRT; 319 AA.
 AC Q99795;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last annotation update)
 DE CELL SURFACE A33 ANTIGEN PRECURSOR.
 OS Homo sapiens (Human).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 CC Eutheria; Primates; Catarrhini; Hominidae; Homo.
 RN [1]
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RC TISSUE-COLON CARCINOMA;
 RX MEDLINE; 97165045.
 RA HEATH J.K., WHITE S.J., JOHNSTONE C.N., CATIMEL B., SIMPSON R.J.,
 RA MORITZ R.L., TU G.-F., JI H., WHITEHEAD R.H., GROENEN L.C.,
 RA SCOTT A.M., RITTER G., COHEN L., WELT S., OLD L.J., NICE E.C.,
 RA BURGESS A.W.;
 RT "The human A33 antigen is a transmembrane glycoprotein and a novel
 RT member of the immunoglobulin superfamily.";
 RL Proc. Natl. Acad. Sci. U.S.A. 94:469-474(1997).
 RN [2]
 RP POST-TRANSLATIONAL MODIFICATIONS.
 RX MEDLINE; 97396159.
 RA MORITZ R.L., JI H., HEATH J.K., WHITE S.J., WELT S., OLD L.J.,
 RA SIMPSON R.J.;
 RT "Characterization of posttranslational modifications of human A33
 RT antigen, a novel palmitoylated surface glycoprotein of human
 RT gastrointestinal epithelium.";
 RL Biochem. Biophys. Res. Commun. 236:682-686(1997).
 CC [1] FUNCTION: MAY PLAY A ROLE IN CELL-CELL RECOGNITION AND SIGNALING.
 CC [1] SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
 CC [1] TISSUE SPECIFICITY: EXPRESSED IN NORMAL GASTROINTESTINAL
 CC EPITHELIUM AND IN 95% OF COLON CANCERS.
 CC [1] PTM: N-GLYCOSYLATED, CONTAINS APPROXIMATELY 8 K OF N-LINKED
 CC CARBOHYDRATE.
 CC [1] PTM: PALMITOYLATED.
 CC [1] SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY. CONTAINS
 CC ONE C2-LIKE AND ONE V-LIKE DOMAINS.

 THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

 EMBL; U79725; AAC50957.1; .
 RSP; P06907; INEU.
 MIM; 602171; .
 PFAM; PF00047; ig. 2.
 KW Immunoglobulin domain; Lipoprotein; Palmitate; Glycoprotein;
 FT Transmembrane; Signal; Antigen.
 SIGNAL 1 21
 CHAIN 22 319
 DOMAIN 22 235
 CELL SURFACE A33 ANTIGEN.
 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 236 256
 FT DOMAIN 257 319
 FT DOMAIN 22 138
 FT DOMAIN 145 229
 FT DOMAIN 258 261
 FT CARBOHYD 112 112
 FT CARBOHYD 200 200
 FT CARBOHYD 223 223
 SQ SEQUENCE 319 AA; 35632 MW; C4F90FC7 CRC32;

Query Match 10.6%; Score 143; DB 1; Length 319;
 Best Local Similarity 36.8%; Pred. No. 8.83e-08;
 Matches 25; Conservative 16; Mismatches 21; Indels 6; Gaps 5;

Db 241 IAVGVVAALIIIGIIYCC-CCRGKDDNTEDKADPNREAYEPPEOLRELSREEREED 299
 QY 97 ITMSVLGSLGTVCCYCCRRKRKPKDSERANREQ-EER-RVROEER-RAEMK 152
 Db 300 DYROEQR 307
 QY 153 S-RHDEIR 159

RESULT 3

ID TRHY_HUMAN STANDARD; PRT; 1898 AA.
 AC Q07283;
 DT 01-OCT-1994 (Rel. 30, Created)
 DT 01-OCT-1994 (Rel. 30, Last sequence update)
 DT 01-FEB-1995 (Rel. 31, Last annotation update)
 DE TRICHOHYALIN.
 GN THH OR TRHY OR THL.
 OS Homo sapiens (Human).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 CC Eutheria; Primates; Catarrhini; Hominidae; Homo.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE; 93280194.
 RA LEE S.-C., KIM I.-G., MAREKOV L.N., O'KEEFE E.J., PARRY D.A.D.,
 RA STEINERT P.M.;
 RT "The structure of human trichohyalin. Potential multiple roles as a
 RT functional EF-hand-like calcium-binding protein, a cornified cell
 RT envelope precursor, and an intermediate filament-associated (cross-
 RT linking) protein.";
 RL J. Biol. Chem. 268:12164-12176(1993).
 RN [2]
 RP SEQUENCE OF 1731-1898 FROM N.A., AND CHARACTERIZATION.
 RX MEDLINE; 93315897.
 RA O'KEEFE E.J., HAMILTON E.H., LEE S.-C., STEINERT P.M.;
 RT "Trichohyalin: a structural protein of hair, tongue, nail, and
 RT epidermis.";
 RL J. Invest. Dermatol. 101:65S-71S(1993).
 CC [1] FUNCTION: INTERMEDIATE FILAMENT-ASSOCIATED PROTEIN THAT ASSOCIATES
 CC IN REGULAR ARRAYS WITH KERATIN INTERMEDIATE FILAMENTS (KIF) OF THE
 CC INNER ROOT SHEATH CELLS OF THE HAIR FOLLICLE AND THE GRANULAR
 CC LAYER OF THE EPIDERMIS. IT LATER BECOMES CROSS-LINKED TO KIF BY
 CC ISODIPEPTIDE BONDS. IT MAY SERVE AS SCAFFOLD PROTEIN, TOGETHER
 CC WITH INVOLUCRIN, IN THE ORGANIZATION OF THE CELL ENVELOPE OR EVEN
 CC ANCHOR THE CELL ENVELOPE TO THE KIF NETWORK. IT MAY BE INVOLVED IN
 CC ITS OWN CALCIUM-DEPENDENT POSTSYNTHETIC PROCESSING DURING TERMINAL
 CC DIFFERENTIATION.
 CC [1] SUBUNIT: MONOMER (PROBABLE).
 CC [1] TISSUE SPECIFICITY: FOUND IN THE HARD KERATINIZING TISSUES SUCH AS
 CC THE INNER ROOT SHEATH (IRS) OF HAIR FOLLICLES AND MEDULLA, AND IN
 CC THE FILIFORM PAPILLAE OF DORSAL TONGUE EPITHELIUM (PROBABLE).
 CC [1] DEVELOPMENTAL STAGE: EXPRESSED DURING LATE DIFFERENTIATION OF
 CC THE EPIDERMIS.
 CC [1] DOMAIN: CONSISTS OF NINE DOMAINS. DOMAIN 1 CONTAINS TWO EF-HAND
 CC CALCIUM-BINDING DOMAINS. DOMAINS 2-4, 6, AND 8 ARE ALMOST
 CC ENTIRELY ALPHA-HELICAL, CONFIGURED AS A SERIES OF PEPTIDE REPEATS
 CC OF VARYING REGULARITY, AND ARE THOUGHT TO FORM A SINGLE-STRANDED
 CC ALPHA-HELICAL ROD STABILIZED BY IONIC INTERACTIONS. DOMAIN 6 IS
 CC THE MOST REGULAR AND MAY BIND KIF DIRECTLY BY IONIC INTERACTIONS.
 CC DOMAINS 5 AND 7 ARE LESS WELL ORGANIZED AND MAY INDUCE POINTING TN

(TM)

Result No.	Score	Query Match	Length	DB ID	Description	Pred. No.
1	1381	57.5	358	4 095210	GENETHONIN 1.	2.14e-238
2	163	6.8	579	3 12596	GLUCOAMYLASE G2 (EC 3.	1.28e-08
3	135	5.6	613	2 059222	ALPHA-AMYLASE (EC 3.2.	1.37e-04
4	134	5.6	1378	5 097405	COLLAGEN PRO ALPHA-CHA	1.89e-04
5	128	5.3	149	5 000484	MINI-COLLAGEN PRECURSO	1.24e-03
6	128	5.3	491	10 081504	F9D12.19 PROTEIN.	1.24e-03
7	124	5.2	295	5 020927	F57B7.3 PROTEIN.	4.23e-03
8	125	5.2	638	3 002296	GLUCOAMYLASE.	3.12e-03
9	123	5.1	546	2 0924N9	BETA-AMYLASE (EC 3.2.1	5.74e-03
10	123	5.1	902	5 016161	PRECOLLAGEN P PRECURSO	5.74e-03
11	122	5.1	1366	6 046392	TYPE I PROCOLLAGEN PRO	7.77e-03
12	122	5.1	1453	11 063079	COLLAGEN ALPHA1 (FRAGM	7.77e-03
13	122	5.1	1460	6 09X577	TYPE I COLLAGEN PRO-AL	7.77e-03
14	122	5.1	1461	4 076045	PRO ALPHA 1(I) COLLAG	7.77e-03
15	120	5.0	150	6 09X724	TYPE II COLLOGEN CYANO	1.42e-02
16	119	5.0	280	5 04173	F58F6.2 PROTEIN.	1.91e-02
17	119	5.0	347	6 09X725	TYPE II COLLOGEN CYANO	1.91e-02
18	119	5.0	684	5 090679	FIBRILLAR COLLAGEN (FR	1.91e-02
19	119	5.0	682	2 030565	CYCLODEXTRIN GLYCOSYL	1.91e-02
20	120	5.0	853	2 053439	PGRS-FAMILY PROTEIN	1.42e-02

RESULT	1
ID	O95210 PRELIMINARY; PRT; 358 AA.
AC	O95210;
DT	01-WAY-1999 (TREMBLrel. 10, Created)
DF	01-WAY-1999 (TREMBLrel. 10, Last sequence update)
DT	01-WAY-1999 (TREMBLrel. 10, Last annotation update)
DE	GENETHONIN 1.
OS	Homo sapiens (Human).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC	Eutheria; Primates; Catarrhini; Homnidae; Homo.
RN	(1)
RP	SEQUENCE FROM N.A.
RC	TISSUE=MUSCLE;
RX	MEDLINE; 96423696.
RA	PIETU G., ALIBERT O., GUICHARD V., LAMY B., BOIS F., LEROY E.,
RA	"MARTAGE-SAMPSON R., HOULGATTE R., SOULARUE P., AUFRAY C.;
RT	Novel gene transcripts preferentially expressed in human muscles
RT	revealed by quantitative hybridization of a high density cDNA array.";
RL	Genome Res. 6:492-503(1996).
RN	[2]
RP	SEQUENCE FROM N.A.
RC	TISSUE=MUSCLE;
RX	MEDLINE; 99013693.
RA	BOUJU S., LIGNON M.-F., PIETU G., LE CUNFF M., LEGER J.-J.,
RA	AUFRAY C., DECESNE C.A.;
RT	"Molecular cloning and functional expression of a novel human gene
RT	encoding two 41-43 kDa skeletal muscle internal membrane proteins.";
RL	Biochem. J. 335:549-556(1998).
DR	EMBL; AF062534; AAC78827.1;
SQ	SEQUENCE 358 AA; 39007 MW; BBC512D6 CRC32;
Query Match 57.5%; Score 1381; DB 4; Length 358;	
Best Local Similarity 62.6%; Pred.No. 2,14e-238;	
Matches 226; Conservative 53; Mismatches 56; Indels 26; Gaps 1	
Db	1 MGAVSALLVGGGLAGALFWLLRGGPDGTGKGDGEQ-EKDAPLGGAATPG-CHQSGS 58 : : : :
Qy	1 MGAVSALLVGGGLAGALILWLRDSGAPGKGVAEPQKGAPPGEAAAPGDPGPGGGS 60
Db	59 G-LSPGPSGOELVTKPEHLQESNGHLSIKTKDLQLQAASWRLQNPSREVCDSNRHVPS 117 : : : :
Qy	61 GGLSPEFSDELYSKAEHLRESNGHLSIESKDLGNLPEAQ-RLQNVGDWV-NAREFPVP 118 : : : :
Db	118 GPPDTEPATSTSNRSYSVEVSERNLESIPMGEGFKQGOETSAAKYCTFAEKLPSN 177

QY 119 GRIPOTHSRADSEAA--RN--Q--SPG--S--HG--GEWRLPKGOETAVKVGASVAAKLASS 169
 Db 178 LKKNRAKEEMSLDLNSQDRVDHEHEKEMVPRHSSWDGVDGSGSLKAPVLNLCQDNGRS 237
 QY 170 LLDVDRAK-----A-V-SQDQAGHEDMEVSRHSSWGSVGLGGSLEASRLSLNQRMDSTN 222
 Db 238 TLVEARGQOVHGKMERVAVMPAGSQVSRVFOVHYTSTDVQFIATVDGHECIGRNTYI 297
 QY 223 SLVGGGWEVDKVA--SLKP---QVQSIFQVHYTTNDVQFIATVDGHECIGRNTYI 277
 Db 298 PLHYNKGDFMWSHIFLPADTVVEMKPVLENKVTWECSNRFLETGHEKDVVHAWGI 357
 QY 278 PLHYCKDGLWSHVSFLPADTVVEMKPVLENKVTWECSNRFLETGHEKDVVHAWGI 337
 Db 358 H 358
 QY 338 H 338

RESULT 2
 ID Q12596 PRELIMINARY; PRT; 579 AA.
 AC Q12596;
 DT 01-NOV-1996 (TREMELrel. 01, Created)
 DT 01-NOV-1996 (TREMELrel. 01, Last sequence update)
 DE GLUCOAMYLASE G2 (EC 3.2.1.3).
 OS Corticium rolfsii.
 OC Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Stereales;
 OC Corticiaceae; Corticium.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-AH9627;
 RX MEDLINE; 96158471.
 RA NAGASAKA Y., MURAKI N., KIMURA A., SUTO M., YOKOTA A., TOMITA F.;
 RT "Cloning of Corticium rolfsii glucoamylase cDNA and its expression in
 RT Saccharomyces cerevisiae".
 RL Appl. Microbiol. Biotechnol. 44:451-458(1995).
 DR EMBL; D49448; BAA08436.1; -;
 DR HSSP; P04064; IGLM.
 DR PFAM; PF00686; CBD_4; 1.
 DR PFAM; PF00723; Glyco_hydro.15; 1.
 DR PRINTS; PR00736; GLHYDRASE15.
 KW Hydrolase; Glycosidase.
 SQ SEQUENCE 579 AA; 61548 MW; 0439FA95 CRC32;

Query Match 6.8%; Score 163; DB 3; Length 579;
 Best Local Similarity 36.3%; Pred. No. 1.28e-08;
 Matches 29; Conservative 16; Mismatches 30; Indels 5; Gaps 5;

Db 483 VEVTFDV-YATVYQNIYITGDVSELGNWTPANGVALSSANYPTWSATIALPADTTIQY 541
 QY 244 VSIOFQVHYTTNDVQFIATVDGHESLGRWN-TY-IPLHYCKDGLWSHVSFLPADTVVEM 301
 Db 542 KYVNDGSTVI-WEDAISNR 560
 QY 302 KFLVLENKVTWECSN-R 320

RESULT 3
 ID Q59222 PRELIMINARY; PRT; 613 AA.
 AC Q59222;
 DT 01-NOV-1996 (TREMELrel. 01, Created)
 DT 01-NOV-1996 (TREMELrel. 01, Last sequence update)
 DE ALPH-A-MYLASE (EC 3.2.2.1) (PURINE NUCLEOSIDASE).
 GN AMY.
 OS Bacillus sp.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group;
 OC Bacillus/Staphylococcus group; Bacillus.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-TS-23;

RA LIN L.L., CHU W.S., HSU W.H.;
 RL Submitted (MAR-1995) to the EMBL/Genbank/DBJ databases.
 CC 1- CATALYTIC ACTIVITY: A N-D-RIBOSYL-PURINE + H(2)O -> A PURINE + D-
 CC RIBOSE.
 DR EMBL; U22045; AAA63900.1; -;
 DR HSSP; P06278; 1BPL.
 DR PFAM; PF00128; alpha-amylase; 1.
 DR PFAM; PF00686; CBD_4; 1.
 KW Hydrolase; Glycosidase.
 SQ SEQUENCE 613 AA; 69536 MW; DC66587C CRC32;

Query Match 5.6%; Score 135; DB 2; Length 613;
 Best Local Similarity 27.7%; Pred. No. 1.37e-04;
 Matches 31; Conservative 17; Mismatches 60; Indels 4; Gaps 4;

Db 483 LTGNRSDTVTINADGGEFKVNGSVSIWAKTSNVTFTVNNATTSGQNVYVANIPEL 542
 QY 213 LNQRMDSTNSLVGGRG-WEVDGKVASLKPQVS-IQFQVHYTTNDVQFIATVDGHESL 270
 Db 543 GNWNTANAKMNPSSYPTWKATIALPOGKAIEPKFKKDOAGNVIWESTSNR 594
 QY 271 GRNNT-Y-IPLHYCKDGLWSHVSFLPADTVVEMKPVLENKVTWECSNR 320

RESULT 4
 ID O97405 PRELIMINARY; PRT; 1378 AA.
 AC O97405;
 DT 01-MAY-1999 (TREMELrel. 10, Created)
 DT 01-MAY-1999 (TREMELrel. 10, Last sequence update)
 DE COLLAGEN PRO ALPHA-CHAIN PRECURSOR.
 GN HDCOL 1 ALPHA.
 OS Haliotis discus (Abalone).
 OC Eukaryota; Metazoa; Mollusca; Gastropoda; Archaeogastropoda;
 OC Haliotidae; Haliotis.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-MUSCLE;
 RA WATABE S.;
 RT "Two types of collagen pro alpha chain in the abalone muscle".
 RL Submitted (SEP-1998) to the EMBL/Genbank/DBJ databases.
 DR EMBL; AB017600; BAA75668.1; -;
 KW Signal; Collagen.
 FT SIGNAL 1 18 POTENTIAL.
 SQ SEQUENCE 1378 AA; 132583 MW; 387BEA97 CRC32;

Query Match 5.6%; Score 134; DB 5; Length 1378;
 Best Local Similarity 40.4%; Pred. No. 1.89e-04;
 Matches 19; Conservative 13; Mismatches 14; Indels 1; Gaps 1;

Db 801 RGSGGIPGNSGTPGDPGRAGPPGSPGAGP-PPGSGATGLSGDGER 846
 QY 24 RDSGAPGKGVNAEPQKGPAGPGEAAAPGDPGGGSGGLSPESDR 70

RESULT 5
 ID Q00484 PRELIMINARY; PRT; 149 AA.
 AC Q00484;
 DT 01-NOV-1996 (TREMELrel. 01, Created)
 DT 01-NOV-1996 (TREMELrel. 01, Last sequence update)
 DE MINI-COLLAGEN PRECURSOR (ISOFORM 1).
 GN N-COL 1.
 OS Hydra attenuata (Hydra) (Hydra vulgaris).
 OC Eukaryota; Metazoa; Cnidaria; Hydrozoa; Hydrozoa; Anthomedusae;
 OC Hydridae; Hydra.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-SF1;
 RX MEDLINE; 92064646.
 RA KURZ E.M., HOLSTEIN T.W., PETRI B.M., ENGEL J., DAVID C.N.;
 RT "Mini-collagens in hydra nematocytes".
 RL J. Cell Biol. 115:1159-1169(1991).

W P S R E L H
(TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MParch_pp protein - protein database search, using Smith-Waterman algorithm
Run on: Wed Apr 12 07:52:24 2000; MasPar time 7.02 Seconds
Tabular output not generated. 401.344 Million cell updates/sec

Title: >US-09-276-268-17
Description: (1-119) from US09276268.pep
Perfect Score: 892
Sequence: 1 GTPASVLRVSVDPSLPP.....LGGFSCQVRLNKRKEYMVR 119

Scoring table: PAM 150
Gap 11

Searched: 189963 seqs, 23686106 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: a-geneseq35
1:geneseqp

Statistics: Mean 29.971; Variance 126.276; scale 0.237

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description	Pred. No.
1	763	85.5	98	1	W85614	3.13e-63
2	420	47.1	56	1	Y13130	1.90e-29
3	292	32.7	41	1	Y12330	3.31e-17
4	93	10.4	482	1	R21409	1.22e-01
5	88	9.9	167	1	W73235	2.91e-01
6	88	9.9	167	1	W83918	2.91e-01
7	87	9.8	336	1	W23397	3.46e-01
8	86	9.6	572	1	R39284	4.10e-01
9	85	9.5	273	1	W20567	4.87e-01
10	85	9.5	300	1	W20098	4.87e-01
11	83	9.3	341	1	F03221	6.82e-01
12	83	9.3	341	1	W01456	6.82e-01
13	83	9.3	463	1	W88315	6.82e-01
14	83	9.3	520	1	R06261	6.82e-01
15	82	9.2	137	1	R07292	8.07e-01
16	81	9.1	93	1	W20762	9.54e-01
17	81	9.1	131	1	W48322	9.54e-01
18	81	9.1	224	1	W97727	9.54e-01
19	81	9.1	334	1	W19854	9.54e-01
20	81	9.1	476	1	R70099	9.54e-01
21	81	9.1	967	1	W93621	9.54e-01
22	80	9.0	31	1	W97645	1.13e-02
23	80	9.0	39	1	Y13175	1.13e-02

24	80	9.0	73	1	W20195	H. pylori membrane pro	1.13e+02
25	80	9.0	97	1	W97643	Breast-specific BS202	1.13e+02
26	80	9.0	229	1	W02737	Human secreted protein	1.13e+02
27	80	9.0	243	1	W80702	S. pneumoniae cation t	1.13e+02
28	80	9.0	256	1	W25958	ORF 2 protein of PRRSV	1.13e+02
29	80	9.0	334	1	W22732	Human ATP receptor	1.13e+02
30	80	9.0	420	1	W10168	Human regulator of G-p	1.13e+02
31	80	9.0	517	1	R06260	Human acetylcholine re	1.13e+02
32	80	9.0	600	1	R95269	Nisin nist gene produc	1.13e+02
33	80	9.0	1194	1	R42199	HSMGluR1	1.13e+02
34	80	9.0	1199	1	R25080	GLU-G-R subtype Ia	1.13e+02
35	80	9.0	1219	1	W25763	Amino acid sequence of	1.13e+02
36	80	9.0	2016	1	W23994	Human hH1 sodium chann	1.13e+02
37	79	8.9	290	1	W98335	H. pylori GHPO 18 prot	1.33e+02
38	79	8.9	290	1	W71474	Helicobacter polypepti	1.33e+02
39	79	8.9	411	1	W97718	Staphylococcus aureus	1.33e+02
40	79	8.9	531	1	R77173	Condensing enzyme clon	1.33e+02
41	79	8.9	816	1	R44143	Rabbit sodium ion/hydr	1.33e+02
42	79	8.9	854	1	W37991	Mutant Aspergillus ory	1.33e+02
43	79	8.9	882	1	R43996	Thermophilic DNA polym	1.33e+02
44	79	8.9	882	1	R45749	Alpha-DNA polymerase	1.33e+02
45	78	8.7	806	1	R65495	Marek's disease virus	1.57e+02

ALIGNMENTS

RESULT 1
ID W85614 standard; Protein; 98 AA.
AC W85614;
DE 02-MAR-1999 (first entry)
KW Clone; secreted protein; protein factor; cytokine; lymphokine;
KW interferon; colony stimulating factor; CSF; interleukin; cloning;
KW tumour invasion; tumour suppression; immune boosting.
OS Homo sapiens.
FH Key Location/Qualifiers
FT Peptide 25..37
FT Peptide /note= "Predicted signal peptide/transmembrane domain"
FT Peptide 62..74
FT Peptide /note= "Possible signal peptide/transmembrane domain"

WO9849302-A1.
05-NOV-1998.
PF 24-APR-1998; U08336.
PR 23-APR-1998; US-065125.
PR 25-APR-1997; US-845296.
PA (GEMV) GENETICS INST INC.
PI Agostino MJ, Jacobs K, Lavallie ER, McCoy JM, Merberg D,
PI Racie LA, Spaulding V, Treacy M;
DR WPI; 99-024059/02.
DR N-PSDB; V83139.
DR New polynucleotides encoding secreted human proteins - are derived
PT from human foetal brain, adult brain, adult blood or placenta CDNA
PT libraries, useful, e.g. as potential immunomodulators
PS Claim 30; Page 81; 104pp; English.
CC The nucleotide sequence (NS) of the full-length protein-coding
CC sequence of clones ci25.4 (V83132), da228.6 (V83133), du410.5
CC (V83134), eh80.1 (V83135), er369.1 (V83136), fh123.5 (V83137),
CC fnd0.1 (V83138) or fr473.2 (V83139), (all clones are deposited as
CC ATCC 98415) and the proteins they encode are predicted to have
CC biological activities which would make them suitable for treating,
CC preventing or ameliorating medical conditions in humans and animals
CC for example, tumour suppression/invasion activity, immune system
CC boosting activity. The polynucleotides are also believed to be useful
CC for gene therapy.
SQ Sequence 98 AA;

Query Match 85.5%; Score 763; DB 1; Length 98;
Best Local Similarity 100.0%; Pred No. 3.13e-63;
Matches 98; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 MASLCCGPKLACAGVLSANGVIMLIMLGIFFNVHSAVLIEDVPFTEKDFENGQMIYN 60

```

QY 22 MASLCCGPKLAACGIVLSANGVIMLMGLGFFNVHSAVLIEDVPFTEKDFENGQNIYN 81
DB 61 LYEQVSYNCFIAGAGLYLLGGFSCQVRNKRKEYMVR 98
QY 82 LYEQVSYNCFIAGAGLYLLGGFSCQVRNKRKEYMVR 119

RESULT 2
ID Y13130 standard; Protein; 56 AA.
AC Y13130;
DE Human secreted protein encoded by 5' EST SEQ ID NO: 144.
KW Human; secreted protein; EST; expressed sequence tag; diagnosis;
KW forensic; gene therapy; chromosome mapping; signal peptide;
KW upstream regulatory sequence; cytokine activity; cell proliferation;
KW differentiation; haematopoiesis regulation; tissue growth regulation;
KW reproductive hormone regulation; chemotactic; chemokinetic; haemostatic;
KW thrombolytic; anti-inflammatory; tumour inhibition.
OS Homo sapiens.
PN WO9906552-A2.
PD 11-FEB-1999.
PF 31-JUL-1998; IB1236.
PR 01-AUG-1997; US-905223.
PA (GEST ) GENSET.
PI Duclert A, Dumas Milne Edwards J, Lacroix B;
DR WPI; 99-153782/13.
DR N-PSDB; X51930.
PT New isolated brain-derived nucleic acids - used to develop products
PT regulating, anti-inflammatory, immune, regulatory, haematopoiesis
PT Claim 34; Page 524; 57pp; English.
PS X51787 to X52019 represent 5' expressed sequence tags (ESTs) for human
CC secreted proteins, and encode the proteins given in Y12987 to Y13219,
CC respectively. The proteins given represent the signal peptide and an
CC N-terminal fragment of a secreted protein. The nucleic acid sequences
CC can be used for producing secreted human gene products. They can also
CC be used to develop products for diagnosis and therapy. The proteins
CC obtained may have cytokine activity, cell proliferation/differentiation
CC activity, haematopoiesis regulating activity, tissue growth regulating
CC activity, reproductive hormone regulating activity, chemotactic/
CC chemokinetic activity, haemostatic and thrombolytic activity, receptor/
CC ligand activity, anti-inflammatory activity, tumour inhibition activity
CC or other activities. The products can be used in forensic, gene therapy
CC and chromosome mapping procedures. The sequences can also be used for
CC obtaining corresponding promoter sequences. The nucleic acids encoding
CC the signal peptide can be used for directing extracellular secretion of
CC a polypeptide or the insertion of a polypeptide into a membrane, or
CC importing a polypeptide into a cell.
SQ Sequence 56 AA;

Query Match 47.1%; Score 420; DB 1; Length 56;
Best Local Similarity 98.2%; Pred. No. 1.90e-29;
Matches 55; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 1 MASLCCGPKLAACGIVLSANGVIMLMGLGFFNVHSAVLIEDVPFTEKDFENGQ 56
QY 22 MASLCCGPKLAACGIVLSANGVIMLMGLGFFNVHSAVLIEDVPFTEKDFENGQ 77

RESULT 3
ID Y12330 standard; Protein; 41 AA.
AC Y12330;
DE Human 5' EST secreted protein SEQ ID NO:361.
KW Human; secreted protein; EST; expressed sequence tag; diagnosis;
KW forensic; gene therapy; chromosome mapping; signal peptide;
KW upstream regulatory sequence; cytokine activity; cell proliferation;
KW differentiation; haematopoiesis regulation; tissue growth regulation;
KW reproductive hormone regulation; chemotactic; chemokinetic; haemostatic;
KW thrombolytic; anti-inflammatory; tumour inhibition.
OS Homo sapiens.
PN WO9906548-A2.

```

```

PD 11-FEB-1999.
PF 31-JUL-1998; IB1222.
PR 01-AUG-1997; US-905135.
PA (GEST ) GENSET.
PI Duclert A, Dumas Milne Edwards J, Lacroix B;
DR WPI; 99-153778/13.
DR N-PSDB; X41163.
PT New nucleic acids encoding human secreted proteins - obtained from
PT cDNA libraries prepared from e.g. liver, ovary, brain, prostate,
PT kidney, lung, umbilical cord, placenta and colon tissue
PS Claim 27; Page 695; 824pp; English.
CC X41094 to X41347 represent 5' expressed sequence tags (ESTs) for human
CC secreted proteins, and encode the proteins given in Y12261 to Y12514,
CC respectively. The proteins given represent the signal peptide and an
CC N-terminal fragment of a secreted protein. The nucleic acid sequences
CC can be used for producing secreted human gene products. They can also
CC be used to develop products for diagnosis and therapy. The proteins
CC obtained may have cytokine activity, cell proliferation/differentiation
CC activity, haematopoiesis regulating activity, tissue growth regulating
CC activity, reproductive hormone regulating activity, chemotactic/
CC chemokinetic activity, haemostatic and thrombolytic activity, receptor/
CC ligand activity, anti-inflammatory activity, tumour inhibition activity
CC or other activities. The products can be used in forensic, gene therapy
CC and chromosome mapping procedures. The sequences can also be used for
CC obtaining corresponding promoter sequences. The nucleic acids encoding
CC the signal peptide can be used for directing extracellular secretion of
CC a polypeptide or the insertion of a polypeptide into a membrane, or
CC importing a polypeptide into a cell.
SQ Sequence 41 AA;

Query Match 32.7%; Score 292; DB 1; Length 41;
Best Local Similarity 95.0%; Pred. No. 3.31e-17;
Matches 38; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Db 1 MLMLGIFNVHSAVLIEDVPFTEKDFENGQNIYNLYEH 40
QY 46 MLMLGIFNVHSAVLIEDVPFTEKDFENGQNIYNLYEQ 85

RESULT 4
ID R21409 standard; Protein; 482 AA.
AC R21409;
DE 01-APR-1992 (first entry)
DE NADH dehydrogenase subunit 2.
KW Pneumonia; assay; AIDS; immunosuppressed.
OS Pneumocystis carinii.
PN WO9119005-A.
PD 12-DEC-1991.
PF 31-MAY-1991; G00869.
PR 01-JUN-1990; GB-012196.
PA (ISIS-) ISIS INNOVATION LTD.
PI Wakefield AE, Hopkin JM, Moxon ER;
DR WPI; 92-007487/01.
DR N-PSDB; Q20065.
PT New DNA sequences which act as oligo-nucleotide primers - for
PT assaying DNA sample from respiratory secretion of a patient
PT infected with P carinii
PS Claim 3; Fig 3; 42pp; English.
CC The amino acid sequence is that of P. carinii NADH dehydrogenase
CC subunit 2 which was translated from DNA from plasmid pAZ112. See
CC also R21410-R21413 and R20056.
SQ Sequence 482 AA;

Query Match 10.4%; Score 93; DB 1; Length 482;
Best Local Similarity 28.2%; Pred. No. 1.22e-01;
Matches 20; Conservative 18; Mismatches 30; Indels 3; Gaps 3;

Db 8 SOLIAIASSSHNVLVLSRISIIILYVYVEI-IGLIG-IGIYNGFIQVTS 65
QY 30 PKLAACGIVLSANGVIMLMGLGFFNVHSAVLIEDVPFTEKDFENGQNIYNLYEQV-SY 88
DB 66 TQVDIFELL 76

```

Wed Apr 12 08:13:57 2000

(TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd.

Mpsrch_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Wed Apr 12 08:04:32 2000; MasPar time 13.26 Seconds

Tabular output not generated.

```

Title:
Description:
Perfect Score:
Sequence:

```

Scoring table: PAM 150
Gap 11

Searched: 188963 seqs, 23686106 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: a-geneseq35
1:geneseqp

Statistics: Mean 33.327; Variance 130.318; scale 0.256

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query		DB	ID	Description	Pred. No.
	Score	Match Length				
1	1575	66.2	263	Y01098	Human follistatin-3 pr	6.11e-155
2	796	33.5	317	R74600	Follistatin.	1.22e-70
3	789	33.2	319	P93396	Human follistatin (hfs	6.81e-70
4	785	33.0	304	R20062	Rat EDF-binding protei	1.82e-65
5	783	32.9	304	R20063	Human EDF-binding prot	2.93e-69
6	778	32.7	304	R20061	P19 EDF-binding protei	1.02e-68
7	510	21.4	462	R70624	Human polypeptide chai	2.54e-40
8	506	21.3	462	W11237	Elongation factor-1alp	6.65e-40
9	496	20.9	460	W03518	Prostate tumour induci	7.40e-39
10	428	18.0	460	W36114	Yeast elongation facto	8.79e-32
11	393	16.5	293	R05567	plant elongation facto	3.59e-28
12	393	16.5	447	R31594	61-1 clone polypeptide	3.59e-28
13	392	16.5	458	R33050	Translation elongation	4.55e-28
14	388	16.3	294	R05569	Plant elongation facto	1.17e-27
15	388	16.3	448	R31593	LeEF-1 protein.	1.17e-27
16	388	16.3	448	R31595	LeEP-1 clone A polypep	1.17e-27
17	388	16.3	448	R05565	Plant elongation facto	1.17e-27
18	188	7.9	664	W04329	Human HP-8 antigen ass	8.96e-08
19	170	7.1	380	W17860	Follistatin related pr	4.28e-06
20	168	7.1	380	W09406	Transforming growth fa	6.55e-06
21	160	6.7	499	R37508	Human GSPT1-TK.	3.55e-05
22	143	6.0	303	R05164	Sequence of human bone	1.21e-03
23	140	5.9	374	W07663	Human transforming gro	2.23e-03

24	140	5.9	374	1	W37497	Human TMP-2.	2.23e-03
25	133	5.6	500	1	W4496	Human delta-2 protein	9.20e-03
26	133	5.6	659	1	W94497	Human delta-2 mature p	9.20e-03
27	133	5.6	685	1	W80813	Nucleotide sequence of	9.20e-03
28	133	5.6	685	1	W94507	Human delta-2 protein.	9.20e-03
29	129	5.4	44	1	W58633	HP-8 auto-antigen, rei	2.05e-02
30	129	5.4	44	1	W04328	Human HP-8 antigen epi	2.05e-02
31	125	5.3	51	1	W04330	Human HP-8 antigen epi	4.54e-02
32	127	5.3	86	1	W62074	Human tumour-associate	3.05e-02
33	127	5.3	86	1	W15067	Human PEC-60-type prot	3.05e-02
34	122	5.1	104	1	R36608	Thrombin inhibitory pr	8.19e-02
35	118	5.0	282	1	R79101	Prostaglandin I2 (PGI2	1.79e-01
36	118	5.0	282	1	R79102	Prostaglandin I2 (PGI2	1.79e-01
37	118	5.0	282	1	X06933	Angiomodulin/tumour-de	1.79e-01
38	118	5.0	548	1	R33741	Xrl.	1.79e-01
39	106	4.5	436	1	W47422	Human secreted protein	1.78e+00
40	106	4.5	873	1	R44735	apo-E lipoprotein rece	1.78e+00
41	108	4.5	1801	1	W50895	Rat laminin B2 chain.	1.22e+00
42	108	4.5	1964	1	W95557	Rat laminin B2 chain.	1.22e+00
43	105	4.4	632	1	W39256	Mus musculus notch4 pr	2.15e+00
44	105	4.4	846	1	R74692	Rat very low density l	2.15e+00
45	105	4.4	1251	1	R79475	Mouse L3SP-3.	2.15e+00

ALIGNMENTS

RESULT 1

ID Y01098 standard; Protein; 263 AA.

AC AC Y01098;

DT 11-JUN-1999 (first entry)

DE Human follistatin-3 protein sequence.

KW Follistatin-3; human; cancer; cellular growth disorder; liver cirrhosis;

KW differentiation disorder; reproductive system disorder; male sterility;

KW activin-induced differentiation inhibitor; gonadotroph adenoma; hepatoma;

KW osteosarcoma; idiopathic pulmonary fibrosis; pulmonary fibrosis; tumour;

KW fibrotic disorder; osteoarthritis; haematopoiesis; infectious disease;

KW sepsis; cancer; silicosis; sarcoidosis; endotoxic shock; therapy.

OS Homo sapiens.

WS W09910364-A1.

PN PN

PD 04-MAR-1999.

PF 27-AUG-1998.

PF 29-AUG-1997; UI7710.

PR (HUMA-) HUMAN GENOME SCI INC.

PA Duan R, Ruben SM;

PI MPI; 99-204546/17.

DR N-PSDB; X28124.

DR New follistatin-3 polypeptides and nucleic acids - used to develop

PT products for treating e.g. cancers, male sterility, wound healing,

PT fibrotic disorders, angiogenesis and autoimmune, inflammatory and

PT infective diseases

PS Claim 18; Fig 1; 109pp; English.

CC This sequence is the follistatin-3 (FS3) protein of the invention..

CC The products can be used to treat cancers and other cellular growth and

CC differentiation disorders as well as disorders of the reproductive

CC system. FS3 can be used or to treat male sterility. FS3 may also be used

CC to inhibit the activin-induced differentiation of follicular granulosa

CC cells. FS3 may be used therapeutically to regulate autocrine endothelial

CC cell activity and, as a result, induce angiogenesis. Treatment to

CC increase the expression or the presence of FS3 may be used to inhibit the

CC progression of gonadotroph adenomas, osteosarcomas, hepatomas, and other

CC tumours and cancers. FS3 may also be used to treat other fibrotic

CC disorders including liver cirrhosis, osteoarthritis and pulmonary

CC fibrosis. It may also be used to regulate haematopoiesis, and to treat

CC sepsis. Antagonists of FS3 may be used to treat a deficiency in FSH,

CC oestrogen and other hormones, to prevent or inhibit or reduce the

CC production of spermatozoa, to modulate gonadal androgen biosynthesis. FS3

CC antagonists may also be used to treat infectious diseases including

CC silicosis, sarcoidosis, idiopathic pulmonary fibrosis by altering the

CC activation state of mononuclear phagocytes, to treat idiopathic

CC hyper-eosinophilic syndrome by preventing eosinophil production and

CC activation. Endotoxic shock may also be treated by FS3 antagonists by

CC preventing the activation of macrophages.

CC Sequence 263 AA;

CC SQ

WPELEH (TM)

Release 3.1A John F. Collins, Blocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPSrch_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Wed Apr 12 08:05:05 2000; MasPar time 8.58 Seconds
Tabular output not generated. 478.795 Million cell updates/sec

Title: >US-09-276-268-20
Description: (1-317) from US09276268.pep
Perfect Score: 2378
Sequence: 1 MRSGALPMLMGALVWTGVS.....KKAAGAGKVTKSQAQKAK 317

Scoring table: PAM 150
Gap 11

Searched: 131253 seqs, 12956647 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: a-issued

Statistics: 1:5A_COMB 2:5B_COMB 3:PCT9_COMB 4:backfiles1

Mean 31.424; Variance 133.180; scale 0.236

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	1575	66.2	263	2	US-08-972- Sequence 2, Applicatio	1.03e-143
2	519	21.8	73	2	US-08-972- Sequence 5, Applicatio	1.41e-38
3	510	21.4	462	4	5225348-1 Patent No. 5225348.	1.06e-37
4	510	21.4	462	2	US-08-371- Sequence 18, Applicati	1.06e-37
5	506	21.3	462	1	US-08-299- Sequence 1, Applicatio	2.58e-37
6	488	20.5	397	2	US-08-371- Sequence 19, Applicati	1.43e-35
7	483	20.3	398	2	US-08-371- Sequence 17, Applicati	4.37e-35
8	453	19.0	71	2	US-08-972- Sequence 4, Applicatio	3.45e-32
9	261	11.0	1940	2	US-08-644- Sequence 30, Applicati	4.97e-14
10	188	7.9	664	1	US-08-421- Sequence 6, Applicatio	1.91e-07
11	160	6.7	499	2	US-08-820- Sequence 40, Applicati	5.18e-05
12	143	6.0	303	4	5340934-13 Patent No. 5340934.	1.42e-03
13	140	5.9	374	2	US-08-820- Sequence 25, Applicati	2.53e-03
14	140	5.9	374	3	PCT-US95-0 Sequence 2, Applicatio	2.53e-03
15	129	5.4	44	1	US-08-421- Sequence 4, Applicatio	2.05e-02
16	125	5.3	51	1	US-08-211- Sequence 1, Applicatio	4.33e-02
17	127	5.3	86	2	US-08-744- Sequence 2, Applicatio	2.98e-02
18	127	5.3	86	2	US-08-149- Sequence 1, Applicatio	2.98e-02
19	122	5.1	103	1	US-08-211- Sequence 3, Applicatio	7.57e-02
20	122	5.1	184	1	US-08-211- Sequence 7, Applicatio	7.57e-02
21	122	5.1	184	1	US-08-211- Sequence 15, Applicati	7.57e-02
22	122	5.1	186	1	US-08-211- Sequence 9, Applicati	7.57e-02
23	122	5.1	368	1	US-08-211- Sequence 17, Applicati	7.57e-02

Query Match

66.2%; Score 1575; DB 2; Length 263;

24	118	5.0	548	1	US-08-463- Sequence 2, Applicatio	1.58e-01
25	118	5.0	548	1	US-08-694- Sequence 2, Applicatio	1.58e-01
26	118	5.0	548	1	US-08-333- Sequence 2, Applicatio	1.58e-01
27	116	4.9	129	1	US-08-211- Sequence 11, Applicati	2.29e-01
28	115	4.8	108	1	US-08-211- Sequence 13, Applicati	1.67e+00
29	105	4.4	1193	2	US-08-400- Sequence 10, Applicati	1.67e+00
30	105	4.4	1251	3	PCT-US95-0 Sequence 3, Applicatio	1.67e+00
31	105	4.4	1252	1	US-08-199- Sequence 3, Applicatio	1.67e+00
32	105	4.4	1252	2	US-08-316- Sequence 3, Applicatio	1.67e+00
33	103	4.3	81	2	US-09-149- Sequence 4, Applicatio	2.37e+00
34	103	4.3	81	2	US-08-744- Sequence 4, Applicatio	2.37e+00
35	102	4.3	1015	1	US-08-537- Sequence 1, Applicatio	2.83e+00
36	102	4.3	2471	1	US-08-185- Sequence 16, Applicati	2.83e+00
37	102	4.3	2471	1	US-08-083- Sequence 19, Applicati	2.83e+00
38	100	4.2	86	2	US-09-149- Sequence 3, Applicatio	4.02e+00
39	100	4.2	86	2	US-08-744- Sequence 3, Applicatio	4.02e+00
40	100	4.2	219	1	US-08-152- Sequence 31, Applicati	4.02e+00
41	100	4.2	219	2	US-08-125- Sequence 18, Applicati	4.02e+00
42	100	4.2	219	2	US-08-460- Sequence 18, Applicati	4.02e+00
43	99	4.2	846	1	US-08-451- Sequence 3, Applicatio	4.79e+00
44	99	4.2	873	1	US-08-393- Sequence 2, Applicatio	4.79e+00
45	99	4.2	1833	3	PCT-US95-0 Sequence 18, Applicati	4.79e+00

ALIGNMENTS

RESULT 1
ID US-08-972-008-2 STANDARD; PRT; 263 AA.
XX
AC xxxxxx
XX
DT
XX
DE
XX
XX

Sequence 2, Application US/08972008

Sequence 2 Application US/08972008
Patent No. 5942420
GENERAL INFORMATION:
APPLICANT: Holtzman, Douglas A.

TITLE OF INVENTION: No. 5942420e1 Molecules of the Follistatin-Related

NUMBER OF SEQUENCES: 5

CORRESPONDENCE ADDRESS:

ADDRESSEE: LAHIVE & COCKFIELD, LLP

STREET: 28 State Street

CITY: Boston

STATE: Massachusetts

COUNTRY: USA

ZIP: 02109

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA: US/08/972,008

FILING DATE:

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Amy E. Mandragouras

REGISTRATION NUMBER: 36,207

REFERENCE/DOCKET NUMBER: MNI-026

TELECOMMUNICATION INFORMATION:

TELEPHONE: (617)227-7400

TELEFAX: (617)742-4214

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 263 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE 263 AA; 27663 MW; 326569 CN;

Best Local Similarity 79.4%; Pred. No. 1.03e-143;

Matches 201; Conservative 27; Mismatches 23; Indels 2; Gaps 2;

Db 6 PGLWPLPWGALAWAGFVSS-MGSGNPAPGGVCMWQOQGEATCSLVLTQDVTREACAS 64

QY 3 SGALWELLWALWVTVGSVAVMGSEDSVPGVCMWQOQGEATCSLVLTQDVTREACAS 62

Db 65 GNIDTAWSLTHPGNKINLIGFLGLVHCLPCKDCDGVCGPKACACMLGGRPR-CBCAP 123

QY 63 GNINTAWSNFTHPGNKISLLGLFLVHCLPCKDCDGVCGPKACACMLGGRPR-CBCAP 122

Db 124 DCSGLPARLVQCGSDGATYRDECELRAACRGHPDLVYRGRCRKSCEHVVCPRPQSCV 183

QY 123 NCBEFFPAGEVQCGSDGATYRDECELRAACRGHPDLVYRGRCRKSCEHVVCPRPQSCV 182

Db 184 VDOTGSAHCVVCAAPCPVPSSPQELCGNNNTYISSCHMRQATCFIAGRSIGVRHAGSC 243

QY 183 VDOTGSAHCVVCAAPCPVPSSPQELCGNNNTYISSCHMRQATCFIAGRSIGVRHAGSC 242

Db 244 AGTPEPPGGESA 256

QY 243 TGGPKFLKSGDAA 255

RESULT 2

ID US-08-972-008-5 STANDARD; PRT; 73 AA.

XX

AC

XX

DT

XX

DE

XX

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

Sequence 5, Application US/08972008

Patent No. 5942420

GENERAL INFORMATION:

APPLICANT: Holtzman, Douglas A.

TITLE OF INVENTION: No. 5942420el Molecules of the Follistatin-Related

NUMBER OF SEQUENCES: 5

CORRESPONDENCE ADDRESS:

ADDRESSEE: LAHIVE & COCKFIELD, LLP

STREET: 28 State Street

CITY: Boston

STATE: Massachusetts

COUNTRY: USA

ZIP: 02109

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/972,008

FILING DATE:

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Amy E. Mandragouras

REGISTRATION NUMBER: 36,207

REFERENCE/DOCKET NUMBER: MNI-026

TELEPHONE: (617)227-7400

TELEFAX: (617)742-4214

INFORMATION FOR SEQ ID NO: 5:

SEQUENCE CHARACTERISTICS:

LENGTH: 73 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: peptide

FRAGMENT TYPE: Internal

SEQUENCE 73 AA; 7674 MW; 26282 CN;

Best Similarity 21.8%; Score 519; DB 2; Length 73;

Match 90.4%; Pred. No. 1.41e-38;

Matches 66; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

Db 1 CEHVVCPRPQSCVVDQTSAGVVCRAAPCPVPSPQELCGNNNTYISSCHMRQATCF 60

QY 170 CAQVVCPRPQSCVVDQTSAGVVCRAAPCPVPSPQELCGNNNTYISSCHMRQATCF 229

Db 61 LGRSIGVRHAGSC 73

QY 230 LGRSIGVRHAGSC 242

RESULT 3

ID 5225348-1 STANDARD; PRT; 501 AA.

XX

AC

XX

DT

XX

DE

XX

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

Query Match 21.4%; Score 510; DB 4; Length 462;

Best Local Similarity 79.0%; Pred. No. 1.06e-37;

Matches 79; Conservative 5; Mismatches 13; Indels 3; Gaps 3;

Db 363 CHTAHIACTACFAELKEKIDRRSGKLEDDGPKFLKSGDAIYDMVPGKPCVFSFSDYPLG 422

QY 221 CHLRQATC-FLGRSGV-RHPGIC-TGGPKFLKSGDAIYDMVPGKPCVFSFSDYPLG 277

Db 423 RFVRDMRQTVAVGVKAVDKKAAGAGKVTSAQAQAK 462

QY 278 RFVRDMRQTVAVGVKAVDKKAAGAGKVTSAQAQAK 317

RESULT 4

ID US-08-371-377-18 STANDARD; PRT; 462 AA.

XX

AC

XX

DT

XX

DE

XX

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

Sequence 18, Application US/08371377

Sequence 18, Application US/08371377

Patent No. 5851764

GENERAL INFORMATION:

APPLICANT: Fisher, Paul B.

APPLICANT: Shen, Ruogian

TITLE OF INVENTION: DEVELOPMENT OF DNA PROBES AND

TITLE OF INVENTION: IMMUNOLOGICAL REAGENTS SPECIFIC FOR CELL SURFACE-EXPRES

TITLE OF INVENTION: MOLECULES AND TRANSFORMATION-ASSOCIATED GENES

NUMBER OF SEQUENCES: 22

CORRESPONDENCE ADDRESS:

ADDRESSEE: Cooper & Dunham

STREET: 1185 Avenue of the Americas

CITY: New York

STATE: New York

COUNTRY: United States of America

ZIP: 10036

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk